

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/737,319

DATE: 07/30/98

TIME: 15:36:29

INPUT SET: S27800.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

ENTERED

1
2
3 (1) General Information:
4
5 (i) APPLICANT: KAJIWARA, Susumu
6 MISAWA, Norihiko
7 KONDO, Keiji
8
9 (ii) TITLE OF INVENTION: A DNA CHAIN USEFUL FOR INCREASING
10 PRODUCTION OF CAROTENOID
11
12 (iii) NUMBER OF SEQUENCES: 8
13
14 (iv) CORRESPONDENCE ADDRESS:
15 (A) ADDRESSEE: FOLEY & LARDNER
16 (B) STREET: 3000 K Street, N.W.
17 (C) CITY: Washington
18 (D) STATE: D.C.
19 (E) COUNTRY: U.S.A.
20 (F) ZIP: 20007-5109
21
22 (v) COMPUTER READABLE FORM:
23 (A) MEDIUM TYPE: Floppy disk
24 (B) COMPUTER: IBM PC compatible
25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
27
28 (vi) CURRENT APPLICATION DATA:
29 (A) APPLICATION NUMBER: US 08/737,319
30 (B) FILING DATE: 12-NOV-1996
31 (C) CLASSIFICATION:
32
33 (vii) PRIOR APPLICATION DATA:
34 (A) APPLICATION NUMBER: PCT/JP96/00574
35 (B) FILING DATE: 08-MAR-1996
36
37 (vii) PRIOR APPLICATION DATA:
38 (A) APPLICATION NUMBER: JP 51234/1995
39 (B) FILING DATE: 10-MAR-1995
40
41 (viii) ATTORNEY/AGENT INFORMATION:
42 (A) NAME: Bent, Stephen A.
43 (B) REGISTRATION NUMBER: 29,768
44 (C) REFERENCE/DOCKET NUMBER: 081356/0111
45
46 (ix) TELECOMMUNICATION INFORMATION:

RAW SEQUENCE LISTING PATENT APPLICATION US/08/737,319

DATE: 07/30/98

TIME: 15:36:30

INPUT SET: S27800.raw

47 (A) TELEPHONE: (202) 672-5300
 48 (B) TELEFAX: (202) 672-5399
 49
 50
 51 (2) INFORMATION FOR SEQ ID NO:1:
 52
 53 (i) SEQUENCE CHARACTERISTICS:
 54 (A) LENGTH: 251 amino acids
 55 (B) TYPE: amino acid
 56 (D) TOPOLOGY: linear
 57
 58 (ii) MOLECULE TYPE: protein
 59
 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
 61
 62 Met Ser Met Pro Asn Ile Val Pro Pro Ala Glu Val Arg Thr Glu Gly
 63 1 5 10 15
 64
 65 Leu Ser Leu Glu Glu Tyr Asp Glu Glu Gln Val Arg Leu Met Glu Glu
 66 20 25 30
 67
 68 Arg Cys Ile Leu Val Asn Pro Asp Asp Val Ala Tyr Gly Glu Ala Ser
 69 35 40 45
 70
 71 Lys Lys Thr Cys His Leu Met Ser Asn Ile Asn Ala Pro Lys Asp Leu
 72 50 55 60
 73
 74 Leu His Arg Ala Phe Ser Val Phe Leu Phe Arg Pro Ser Asp Gly Ala
 75 65 70 75 80
 76
 77 Leu Leu Leu Gln Arg Arg Ala Asp Glu Lys Ile Thr Phe Pro Gly Met
 78 85 90 95
 79
 80 Trp Thr Asn Thr Cys Cys Ser His Pro Leu Ser Ile Lys Gly Glu Val
 81 100 105 110
 82
 83 Glu Glu Glu Asn Gln Ile Gly Val Arg Arg Ala Ala Ser Arg Lys Leu
 84 115 120 125
 85
 86 Glu His Glu Leu Gly Val Pro Thr Ser Ser Thr Pro Pro Asp Ser Phe
 87 130 135 140
 88
 89 Thr Tyr Leu Thr Arg Ile His Tyr Leu Ala Pro Ser Asp Gly Leu Trp
 90 145 150 155 160
 91
 92 Gly Glu His Glu Ile Asp Tyr Ile Leu Phe Ser Thr Thr Pro Thr Glu
 93 165 170 175
 94
 95 His Thr Gly Asn Pro Asn Glu Val Ser Asp Thr Arg Tyr Val Thr Lys
 96 180 185 190
 97
 98 Pro Glu Leu Gln Ala Met Phe Glu Asp Glu Ser Asn Ser Phe Thr Pro
 99 195 200 205

RAW SEQUENCE LISTING PATENT APPLICATION US/08/737,319

DATE: 07/30/98

TIME: 15:36:31

INPUT SET: S27800.raw

100
 101 Trp Phe Lys Leu Ile Ala Arg Asp Phe Leu Phe Gly Trp Trp Asp Gln
 102 210 215 220
 103
 104 Leu Leu Ala Arg Arg Asn Glu Lys Gly Glu Val Asp Ala Lys Ser Leu
 105 225 230 235 240
 106
 107 Glu Asp Leu Ser Asp Asn Lys Val Trp Lys Met
 108 245 250
 109
 110
 111

112 (2) INFORMATION FOR SEQ ID NO:2:

113 (i) SEQUENCE CHARACTERISTICS:

114 (A) LENGTH: 259 amino acids

115 (B) TYPE: amino acid

116 (D) TOPOLOGY: linear

117

118 (ii) MOLECULE TYPE: protein

119

120 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

121

122 Met Gln Leu Leu Ala Glu Asp Arg Thr Asp His Met Arg Gly Ala Ser

123 1 5 10 15

124

125 Thr Trp Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys Asp Glu

126 20 25 30

127

128 Cys Ile Leu Val Asp Ala Asp Asp Asn Ile Thr Gly His Val Ser Lys

129 35 40 45

130

131 Leu Glu Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu Leu His

132 50 55 60

133

134 Arg Ala Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu Leu Leu

135 65 70 75 80

136

137 Gln Gln Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp Thr Asn

138 85 90 95

139

140 Thr Cys Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu Val Asp

141 100 105 110

142

143 Gln Leu Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys Ala Ala

144 115 120 125

145

146 Ala Ile Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His Gln Leu

147 130 135 140

148

149 Pro Ala Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys Ala Ala

150 145 150 155 160

151

152

RAW SEQUENCE LISTING PATENT APPLICATION US/08/737,319

DATE: 07/30/98
TIME: 15:36:32

INPUT SET: S27800.raw

```

153  Asp Val Gln Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu His Glu
154                      165                      170                      175
155
156  Met Asp Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala Pro Asn
157                      180                      185                      190
158
159  Pro Asp Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu Leu Arg
160                      195                      200                      205
161
162  Gln Met Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp Phe Arg
163                      210                      215                      220
164
165  Ile Ile Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu Asp Ala
166                      225                      230                      235                      240
167
168  Ala Leu Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His His Ile
169                      245                      250                      255
170
171  Asn Glu Ala
172
173
174
175  (2) INFORMATION FOR SEQ ID NO:3:
176
177      (i) SEQUENCE CHARACTERISTICS:
178          (A) LENGTH: 288 amino acids
179          (B) TYPE: amino acid
180          (D) TOPOLOGY: linear
181
182      (ii) MOLECULE TYPE: protein
183
184      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
185
186  Met Thr Ala Asp Asn Asn Ser Met Pro His Gly Ala Val Ser Ser Tyr
187      1                      5                      10                      15
188
189  Ala Lys Leu Val Gln Asn Gln Thr Pro Glu Asp Ile Leu Glu Glu Phe
190                      20                      25                      30
191
192  Pro Glu Ile Ile Pro Leu Gln Gln Arg Pro Asn Thr Arg Ser Ser Glu
193                      35                      40                      45
194
195  Thr Ser Asn Asp Glu Ser Gly Glu Thr Cys Phe Ser Gly His Asp Glu
196                      50                      55                      60
197
198  Glu Gln Ile Lys Leu Met Asn Glu Asn Cys Ile Val Leu Asp Trp Asp
199                      65                      70                      75                      80
200
201  Asp Asn Ala Ile Gly Ala Gly Thr Lys Lys Val Cys His Leu Met Glu
202                      85                      90                      95
203
204  Asn Ile Glu Lys Gly Leu Leu His Arg Ala Phe Ser Val Phe Ile Phe
205                      100                      105                      110

```

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/737,319DATE: 07/30/98
TIME: 15:36:33

INPUT SET: S27800.raw

206
207 Asn Glu Gln Gly Glu Leu Leu Leu Gln Gln Arg Ala Thr Glu Lys Ile
208 115 120 125
209
210 Thr Phe Pro Asp Leu Trp Thr Asn Thr Cys Cys Ser His Pro Leu Cys
211 130 135 140
212
213 Ile Asp Asp Glu Leu Gly Leu Lys Gly Lys Leu Asp Asp Lys Ile Lys
214 145 150 155 160
215
216 Gly Ala Ile Thr Ala Ala Val Arg Lys Leu Asp His Glu Leu Gly Ile
217 165 170 175
218
219 Pro Glu Asp Glu Thr Lys Thr Arg Gly Lys Phe His Phe Leu Asn Arg
220 180 185 190
221
222 Ile His Tyr Met Ala Pro Ser Asn Glu Pro Trp Gly Glu His Glu Ile
223 195 200 205
224
225 Asp Tyr Ile Leu Phe Tyr Lys Ile Asn Ala Lys Glu Asn Leu Thr Val
226 210 215 220
227
228 Asn Pro Asn Val Asn Glu Val Arg Asp Phe Lys Trp Val Ser Pro Asn
229 225 230 235 240
230
231 Asp Leu Lys Thr Met Phe Ala Asp Pro Ser Tyr Lys Phe Thr Pro Trp
232 245 250 255
233
234 Phe Lys Ile Ile Cys Glu Asn Tyr Leu Phe Asn Trp Trp Glu Gln Leu
235 260 265 270
236
237 Asp Asp Leu Ser Glu Val Glu Asn Asp Arg Gln Ile His Arg Met Leu
238 275 280 285
239
240
241

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1099 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 99..851

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/08/737,319

DATE: 07/30/98
TIME: 15:36:33

INPUT SET: S27800.raw

Line

Error

Original Text